

ECONOMETRIA I - RESUMO

Guilherme Argentieri Pastore

May 2025

1 Modelo Econométrico Linear

1.1 Matriz de variância positiva semi-definida e simétrica

Demonstração (psd):

Suponha $a \in \mathbb{R}^d$ sendo um vetor coluna. Sabemos que $\mathbb{V}[Z] = \mathbb{E}[(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])']$, logo:

$$a' \mathbb{V}[Z] a = a' \mathbb{E}[(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])'] a = \mathbb{E}[a'(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])' a]$$

Note que a' possui dimensão $1 \times d$ e $(Z - \mathbb{E}[Z])$ possui dimensão $d \times 1$, portanto, $a'(Z - \mathbb{E}[Z])$ é um escalar. Analogamente, $(Z - \mathbb{E}[Z])' a$ também é um escalar.

$$\therefore a' \mathbb{V}[Z] a = \mathbb{E}[a'(Z - \mathbb{E}[Z])(Z - \mathbb{E}[Z])' a] = \mathbb{E}[\underbrace{(a'(Z - \mathbb{E}[Z]))^2}_{\geq 0}] \geq 0$$

Demonstração (simétrica):

$$\mathbb{V}[Z] = \begin{bmatrix} \mathbb{E}[Z_1^2] - \mathbb{E}^2[Z_1] & \cdots & \mathbb{E}[Z_1 Z_d] - \mathbb{E}[Z_1] \mathbb{E}[Z_d] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[Z_d Z_1] - \mathbb{E}[Z_d] \mathbb{E}[Z_1] & \cdots & \mathbb{E}[Z_d^2] - \mathbb{E}^2[Z_d] \end{bmatrix}$$

Não é difícil ver que $Z_i Z_j = Z_j Z_i \ \forall i \neq j \implies \mathbb{E}[Z_i Z_j] - \mathbb{E}[Z_i] \mathbb{E}[Z_j] = \mathbb{E}[Z_j Z_i] - \mathbb{E}[Z_j] \mathbb{E}[Z_i] \ \forall i \neq j$, isso vale pois a esperança é um operador linear. Logo, a matriz $\mathbb{V}[Z]$ é simétrica.

1.2 Melhor preditor linear da forma $h_*(x) = \alpha_* + \delta'_* x$ se $\mathbb{V}[Z]$ tem posto cheio

Demonstração:

Temos o seguinte problema de predição: $\min \mathbb{E}[(Y - \alpha - \delta'X)^2]$

$$\begin{aligned} [Y - (\alpha + \delta'X)]^2 &= Y^2 - 2Y(\alpha + \delta'X) + \alpha^2 + 2\alpha\delta'X + \delta'XX'\delta \\ &= Y^2 - 2\alpha Y - 2\delta'XY + \alpha^2 + 2\alpha\delta'X + \delta'XX'\delta \\ \therefore \mathbb{E}[(Y - \alpha - \delta'X)^2] &= \mathbb{E}[Y^2] + \alpha^2 - 2\delta'\mathbb{E}[XY] - 2\alpha\mathbb{E}[Y] + 2\alpha\delta'\mathbb{E}[X] + \delta'\mathbb{E}[XX']\delta \end{aligned}$$

Obtendo as CPO's:

$$\begin{aligned} \frac{\partial EQM}{\partial \alpha} &= 2\alpha - 2\mathbb{E}[Y] + 2\delta'\mathbb{E}[X] = 0 \implies \alpha = \mathbb{E}[Y] - \delta'\mathbb{E}[X] \\ \frac{\partial EQM}{\partial \delta} &= -2\mathbb{E}[XY] + 2\alpha\mathbb{E}[X] + 2\mathbb{E}[XX']\delta = 0 \implies \mathbb{E}[XY] = \alpha\mathbb{E}[X] + \mathbb{E}[XX']\delta \end{aligned}$$

Substituindo α na CPO em relação a δ :

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[X']\delta + \mathbb{E}[XX']\delta \implies Cov(X, Y) = \mathbb{V}[X]\delta$$

Note que $rank(\mathbb{V}[X]) = k \implies \exists \mathbb{V}[X]^{-1} \implies Cov(X, Y) = \mathbb{V}[X]\delta$ tem solução única.

$$\therefore \delta_* = \mathbb{V}[X]^{-1}Cov(X, Y) \text{ e } \alpha_* = \mathbb{E}[Y] - \delta'_*\mathbb{E}[X]$$

1.3 O melhor preditor linear também é a melhor aproximação linear de $\mathbb{E}[Y|X]$

Primeiramente, vamos definir a classe de funções lineares assim como a esperança condicional e o melhor preditor linear respectivamente:

$$\begin{aligned} \mathcal{H}_{Lin} &:= \{h(X) = \alpha + X\delta : \alpha \in \mathbb{R}, \delta \in \mathbb{R}^k\} \\ h^*(X) &:= \mathbb{E}[X|Y] \\ h_*(X) &:= \alpha_* + \delta'_*X \end{aligned}$$

$$\text{Devo provar que } \mathbb{E}[(h^*(X) - h_*(X))^2] = \min_{h \in \mathcal{H}_{Lin}} \mathbb{E}[(h^*(X) - h(X))^2]$$

Demonstração:

$$\begin{aligned} Y - h(X) &= (Y - h^*(X)) + (h^*(X) - h(X)) \implies (Y - h(X))^2 = [(Y - h^*(X)) + (h^*(X) - h(X))]^2 \\ \implies \mathbb{E}[(Y - h(X))^2] &= \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h(X))^2] + 2\mathbb{E}[(Y - h^*(X))(h^*(X) - h(X))] \end{aligned}$$

É válido lembrar que $\mathbb{E}[Y - h^*(X)|X] = \mathbb{E}[Y|X] - \mathbb{E}[Y|X] = 0$ e que pela LEI, vale $\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$. Com isso, temos:

$$\begin{aligned}\mathbb{E}[(Y - h^*(X))(h^*(X) - h(X))] &= \mathbb{E}[\mathbb{E}[(Y - h^*(X))(h^*(X) - h(X))|X]] \\ \mathbb{E}[(h^*(X) - h(X))\mathbb{E}[(Y - h^*(X))|X]] &= \mathbb{E}[(h^*(X) - h(X)) \cdot 0] = \mathbb{E}[0] = 0\end{aligned}$$

Note que:

$$\begin{aligned}\mathbb{E}[(Y - h(X))^2] &= \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h(X))^2] + 2 \underbrace{\mathbb{E}[(Y - h^*(X))(h^*(X) - h(X))]}_{=0} \\ \therefore \mathbb{E}[(Y - h(X))^2] &= \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h(X))^2]\end{aligned}$$

Dado que $h_*(X)$ é o melhor preditor linear, entre todas as funções lineares $h \in \mathcal{H}_{\text{Lin}}$, h_* minimiza o EQM, ou seja:

$$\mathbb{E}[(Y - h_*(X))^2] \leq \mathbb{E}[(Y - h(X))^2] \quad \forall h \in \mathcal{H}_{\text{Lin}}$$

Vamos analisar a equação $\mathbb{E}[(Y - h(X))^2] = \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h(X))^2]$ em dois casos.

Caso 1 ($h = h_*$):

$$\mathbb{E}[(Y - h_*(X))^2] = \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h_*(X))^2]$$

Caso 2 (h é qualquer $h \in \mathcal{H}_{\text{Lin}}$):

$$\mathbb{E}[(Y - h(X))^2] = \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h(X))^2]$$

Note que, como $\mathbb{E}[(Y - h_*(X))^2] \leq \mathbb{E}[(Y - h(X))^2] \quad \forall h \in \mathcal{H}_{\text{Lin}}$, temos:

$$\begin{aligned}\mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h_*(X))^2] &\leq \mathbb{E}[(Y - h^*(X))^2] + \mathbb{E}[(h^*(X) - h(X))^2] \\ \therefore \mathbb{E}[(h^*(X) - h_*(X))^2] &\leq \mathbb{E}[(h^*(X) - h(X))^2] \quad \forall h \in \mathcal{H}_{\text{Lin}}\end{aligned}$$

Com isso, conclui-se que $\mathbb{E}[(h^*(X) - h_*(X))^2] = \min_{h \in \mathcal{H}_{\text{Lin}}} \mathbb{E}[(h^*(X) - h(X))^2]$

2 MQO

Primeiramente, considere o seguinte modelo: $Y = X\beta + \epsilon$

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} X'_1 \\ \vdots \\ X'_n \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

*Obs: $X_i \in \mathbb{R}^k \quad \forall i = 1, \dots, n$

Problema de minimização: $\min S(b) = \|Y - Xb\|^2 = (Y - Xb)(Y - Xb)' = Y'Y - 2Y'Xb + b'X'Xb$

CPO:

$$\frac{\partial S}{\partial b} = -2X'Y + 2X'Xb = 0 \implies X'Xb = X'Y$$

2.1 H1 - posto cheio

Devo provar que $\text{rank}(X) = k \implies \exists (X'X)^{-1} \implies \hat{b} = (X'X)^{-1}X'Y$

Demonstração:

Seja $v \in \mathbb{R}^k$ um vetor qualquer.

$$v'XX'v = (Xv)'(Xv) = \|Xv\|^2 \geq 0$$

Como visto no curso de verão, sabemos que $\|Xv\| = 0 \iff Xv = 0$, logo, $\|Xv\|^2 = 0 \iff \|Xv\| = 0$. Por hipótese, a matriz X tem posto cheio, o que significa que $Xv = 0 \iff v = 0$, já que o sistema possui solução única trivial, o que implica que a dimensão do núcleo dessa matriz é 0. Portanto, se $v \neq 0$, necessariamente $\|Xv\| \neq 0 \implies \|Xv\|^2 > 0 \implies v'XX'v > 0$. Se $X'X$ é positiva definida, isso significa que seus autovalores são todos estritamente maiores que 0, o que implica em determinante positivo, logo, a matriz é invertível. Sabendo disso, podemos retornar à CPO e obter:

$$\hat{b} = (X'X)^{-1}X'Y$$

2.2 H2 - exogeneidade

Primeiramente, vamos provar que se $\mathbb{E}[\epsilon|X] = 0$, então \hat{b} é não viesado.

Demonstração:

$$\begin{aligned}
\mathbb{E}[\hat{b}|X] &= \mathbb{E}[(X'X)^{-1}X'Y|X] = (X'X)^{-1}X'\mathbb{E}[Y|X] \\
&= (X'X)^{-1}X'\mathbb{E}[X\beta + \epsilon|X] = \underbrace{(X'X)^{-1}X'X}_I \beta + (X'X)^{-1}X' \underbrace{\mathbb{E}[\epsilon|X]}_{=0} = \beta
\end{aligned}$$

Agora, vamos provar que em um modelo da forma $Y = X_1\beta_1 + X_2\beta_2 + \epsilon$ com $\mathbb{E}[\epsilon|X_1, X_2] = 0$, o estimador associado a variável X_1 será viesado caso X_2 seja omitida (havendo ausência de viés em duas execuções).

Demonstração:

$$\begin{aligned}
\tilde{b}_1 &= (X'_1X_1)^{-1}X'_1Y = (X'_1X_1)^{-1}X'_1(X_1\beta_1 + X_2\beta_2 + \epsilon) \\
&= \underbrace{(X'_1X_1)^{-1}X'_1X_1}_I \beta_1 + \underbrace{(X'_1X_1)^{-1}X'_1X_2}_{:=\hat{\gamma}} \beta_2 + (X'_1X_1)^{-1}X'_1\epsilon \\
&= \beta_1 + \hat{\gamma}\beta_2 + (X'_1X_1)^{-1}X'_1\epsilon \\
\implies \mathbb{E}[\tilde{b}_1|X_1] &= \beta_1 + \mathbb{E}[\hat{\gamma}|X_1]\beta_2 + (X'_1X_1)^{-1}X'_1 \underbrace{\mathbb{E}[\epsilon|X_1]}_{=0} \\
\therefore \mathbb{E}[\tilde{b}_1|X_1] &= \beta_1 + \underbrace{\mathbb{E}[\hat{\gamma}|X_1]\beta_2}_{\text{Viés}}
\end{aligned}$$

Note que $\hat{\gamma}$ é o coeficiente angular da regressão de X_2 em X_1 , logo, não é difícil ver que quando X_1 e X_2 são não-correlacionadas, o estimador \tilde{b}_1 é não viesado. Além disso, também é possível perceber que não haverá viés quando $\beta_2 = 0$.

2.3 H3 - homocedasticidade

A suposição de homocedasticidade impõe que $\exists \sigma^2 > 0$ tal que $\mathbb{V}[\epsilon|X] = \sigma^2 I^2$

2.4 Valores estimados e resíduos

2.4.1 Matriz de projeção

Considere $Y = X\beta + \epsilon$. Sabemos que o estimador de MQO é $\hat{\beta} = (X'X)^{-1}X'Y$, além disso, $\hat{Y} = X\hat{\beta}$, portanto:

$$\hat{Y} = X\hat{\beta} = \underbrace{X(X'X)^{-1}X'}_{:=P} Y = PY$$

P é definida como a matriz de projeção, tendo como propriedades ser simétrica, idempotente e $\text{tr}(P) = \text{rank}(P) = k$.

2.4.2 Matriz residualizadora

$$\hat{\epsilon} = Y - \hat{Y} = Y - PY = \underbrace{(I - P)}_{:=M} Y = MY$$

M é definida como a matriz residualizadora, tendo como propriedades ser simétrica, idempotente, $tr(M) = rank(M) = n - k$ e $X'M = MX = PM = MP = 0$

2.5 Teorema de Gauss-Markov

Sob H1, H2 e H3, o estimador de MQO é **BLUE**, i.e. o não viesado de menor variância na classe dos estimadores lineares.

Demonstração:

Primeiramente tome um estimador linear qualquer $\tilde{\beta} := A'Y$ em que A' é uma matriz em função apenas de X . Podemos ver que:

$$\begin{aligned}\tilde{\beta} &= A'(X\beta + \epsilon) = A'X\beta + A'\epsilon \\ \implies \mathbb{E}[\tilde{\beta}|X] &= \mathbb{E}[A'X\beta|X] + \mathbb{E}[A'\epsilon|X] \\ &= A'X\beta + A' \underbrace{\mathbb{E}[\epsilon|X]}_{=0} = A'X\beta\end{aligned}$$

Para que $\tilde{\beta}$ seja não viesado, é necessário que $A'X = I$, portanto, vamos impor essa restrição e reescrever o estimador, garantindo que ele seja não só linear como também não viesado:

$$\tilde{\beta} = \beta + A'\epsilon \implies \tilde{\beta} - \beta = A'\epsilon \implies \tilde{\beta} - \mathbb{E}[\tilde{\beta}|X] = A'\epsilon$$

Sabendo disso, obtemos a variância condicional de $\tilde{\beta}$:

$$\begin{aligned}\mathbb{V}[\tilde{\beta}|X] &= \mathbb{E}[(\tilde{\beta} - \mathbb{E}[\tilde{\beta}|X])(\tilde{\beta} - \mathbb{E}[\tilde{\beta}|X])'|X] = \mathbb{E}[(\tilde{\beta} - \mathbb{E}[\mathbb{E}[\tilde{\beta}|X]])](\tilde{\beta} - \mathbb{E}[\mathbb{E}[\tilde{\beta}|X]])'|X] \\ &= \mathbb{E}[(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'|X] = \mathbb{E}[(A'\epsilon)(A'\epsilon)'|X] = \mathbb{E}[(A'\epsilon)(\epsilon'A)|X] = A'\mathbb{E}[\epsilon\epsilon'|X]A = \sigma^2 A'A\end{aligned}$$

Agora, vamos obter a variância condicional do estimador de MQO:

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y = (X'X)^{-1}X'(X\beta + \epsilon) = \underbrace{(X'X)^{-1}X'X}_I \beta + (X'X)^{-1}X'\epsilon = \beta + (X'X)^{-1}X'\epsilon \\
\implies \mathbb{V}[\hat{\beta}|X] &= \mathbb{V}[\beta + (X'X)^{-1}X'\epsilon|X] = (X'X)^{-1}X'\mathbb{V}[\epsilon|X]X(X'X)^{-1} \\
&= \sigma^2 \underbrace{(X'X)^{-1}X'X(X'X)^{-1}}_I = \sigma^2(X'X)^{-1}
\end{aligned}$$

Agora, é possível perceber que:

$$\begin{aligned}
\mathbb{V}[\tilde{\beta}|X] - \mathbb{V}[\hat{\beta}|X] &= \sigma^2 A'A - \sigma^2(X'X)^{-1} = \sigma^2[A'A - (X'X)^{-1}] \\
&= \sigma^2[A'A - \underbrace{A'X(X'X)^{-1}X'A}_I] = \sigma^2[A'A - A'PA] = \sigma^2 A'[A - PA] \\
&= \sigma^2 A'[(I - P)A] = \sigma^2 A' \underbrace{(I - P)}_M A = \sigma^2 A'MA \geq 0 \\
\therefore \mathbb{V}[\tilde{\beta}|X] - \mathbb{V}[\hat{\beta}|X] &\geq 0
\end{aligned}$$

Como M é idempotente, por definição, ela também é positiva semi-definida, implicando $A'MA$ também positiva semi-definida.

2.6 H4 - normalidade

Sob esta suposição, vale que $\epsilon|X \sim N(0, \sigma^2 I)$ para amostras finitas.

2.7 Distribuição assintótica

Existem três suposições assintóticas que garantirão a convergência do estimador de MQO, que são:

1. $(X_i, Y_i) \stackrel{i.i.d}{\sim}_k (X, Y)$ com $\mathbb{E}[|Y|^2] < \infty$, $tr(\mathbb{E}[XX']) < \infty$ e $rank(\mathbb{E}[XX']) = k$
2. $\mathbb{E}[X'\epsilon] = 0$
3. X e ϵ possuem quartos momentos finitos

Com essas hipóteses e com H1 e H2 vistas anteriormente, garantimos $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \mathbb{E}[XX']^{-1} \mathbb{E}[XX'\epsilon^2] \mathbb{E}[XX']^{-1})$

Demonstração:

Primeiramente, podemos obter a seguinte expressão para $(\hat{\beta} - \beta)$:

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'(X\beta + \epsilon) \\
&= \underbrace{(X'X)^{-1}X'X}_I \beta + (X'X)^{-1}X'\epsilon \\
&= \beta + (X'X)^{-1}X'\epsilon \\
\therefore (\hat{\beta} - \beta) &= (X'X)^{-1}X'\epsilon
\end{aligned}$$

Com isso, temos que:

$$\begin{aligned}
\sqrt{n}(\hat{\beta} - \beta) &= \sqrt{n}(X'X)^{-1}X'\epsilon \\
&= \frac{n}{n}\sqrt{n}\left(\sum_{i=1}^n X_i X_i'\right)^{-1}\left(\sum_{i=1}^n X_i \epsilon_i\right) \\
&= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n X_i \epsilon_i\right) \\
&= \left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i \epsilon_i\right)
\end{aligned}$$

Note que, pela LGN, sabemos que:

$$\left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right) \xrightarrow{p} \mathbb{E}[XX']$$

Logo, pelo teorema do mapa contínuo, temos:

$$\left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right)^{-1} \xrightarrow{p} \mathbb{E}[XX']^{-1}$$

Ademais, sabendo que $\mathbb{E}[X\epsilon] = 0$, pelo TLC temos:

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i \epsilon_i\right) \xrightarrow{d} \mathcal{N}(0, \mathbb{E}[XX'\epsilon^2])$$

Pelo teorema de Slutsky, como $\left(\frac{1}{n}\sum_{i=1}^n X_i X_i'\right)$ converge em probabilidade e $\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i \epsilon_i\right)$ converge em distribuição, então:

$$\begin{aligned}\sqrt{n}(\hat{\beta} - \beta) &\xrightarrow{d} \mathbb{E}[XX']^{-1} \mathcal{N}(0, \mathbb{E}[XX'\epsilon^2]) \\ \therefore \sqrt{n}(\hat{\beta} - \beta) &\xrightarrow{d} \mathcal{N}(0, \mathbb{E}[XX']^{-1} \mathbb{E}[XX'\epsilon^2] \mathbb{E}[XX']^{-1})\end{aligned}$$

2.8 Consistência do MQO

É interessante notar que para garantir consistência, não precisamos da H2 por completo, precisamos apenas que valha $\mathbb{E}[X'\epsilon] = 0$

Demonstração:

Anteriormente, vimos que $\hat{\beta} = \beta + (X'X)^{-1}X'\epsilon = \beta + \left(\frac{1}{n}X'X\right)^{-1} \left(\frac{1}{n}X'\epsilon\right)$. Pela LGN sabemos que:

$$(X'X) = \left(\sum X_i X_i'\right) \xrightarrow{p} \mathbb{E}[XX']$$

Portanto, pelo teorema do mapa contínuo:

$$\left(\frac{1}{n}X'X\right)^{-1} = \left(\frac{1}{n}\sum X_i X_i'\right)^{-1} \xrightarrow{p} \mathbb{E}[XX']^{-1}$$

Além disso, também pela LGN, sabemos que:

$$\left(\frac{1}{n}X'\epsilon\right) = \left(\frac{1}{n}\sum X_i \epsilon_i\right) \xrightarrow{p} \mathbb{E}[X'\epsilon] = 0$$

Portanto, é possível notar que:

$$\hat{\beta} = \beta + \left(\frac{1}{n}X'X\right)^{-1} \left(\frac{1}{n}X'\epsilon\right) \xrightarrow{p} \beta + \mathbb{E}[XX']^{-1} \underbrace{\mathbb{E}[X'\epsilon]}_{=0} = \beta$$